PROMOTER OF THIOREDOXINE TaTrxh2 IN WHEAT

The invention relates to the cloning and to the characterization of a wheat thioredoxin promoter.

Thioredoxins are low molecular weight proteins which have been demonstrated in a large number of organisms, in which they catalyze various redox reactions involving dithiolsulphhydryl exchanges. Their catalytic site comprises the conserved sequence: -Trp-Cys-Gly/Pro/Ala-Pro-Cys-. Thioredoxins in oxidized form comprise a disulphide bridge, the reduction of which to -SH groups, by reduced ferredoxin or by NADPH, is catalyzed via a specific system.

In plants, 3 types of thioredoxin have been demonstrated: the first 2 (thioredoxins m and f) are ferredoxin-dependant thioredoxins located in the chloroplasts, where they are involved in the regulation of photosynthesis. A third type, named thioredoxin h, has been demonstrated in the cytosol. Thioredoxin h is part of an NADP-dependent thioredoxin system (NTS), in which it is associated with NADPH and with an enzyme named NADP-thioredoxin reductase (NTR).

Initially, 2 thioredoxin h were extracted and partially purified from wheat grain (VOGT and FOLLMANN, Biochem. Biophys. Acta 873, 415-418, 1986). Recently, the inventors' team has isolated and characterized 2 cDNA clones encoding a soft wheat thioredoxin h (TaTrxh1) and a hard wheat thioredoxin h (TdTrxh1) (GAUTIER *et al.*, Eur. J. Biochem. 252, 314-324, 1998). The primary structures deduced from the cDNA clones of the thioredoxins h TaTrxh1 and TdTrxh1 are very conserved (96% identity between them). They have an N-terminal extension which is very rich in Ala residues, the analysis of which reveals a putative transmembrane domain of 20 residues. They show strong homologies with cereal thioredoxin h (70-80%) and dicotyledon thioredoxin h (60%).

The thioredoxin h are involved during germination of the wheat grain, in which they contribute, in the albumen, to mobilizing the stores required for the growth of the embryo. They act in particular:

- by reducing the disulphide bridges of certain storage proteins, such as gliadins and glutenins (KOBREHEL *et al.*, Plant Physiol. 99, 919-924, 1992), which increases their sensitivity to proteolysis;
- by reducing enzymes involved in mobilizing the stores, or inhibitors of these enzymes, which leads to activation of the first and deactivation of the second.

The use of thioredoxin h has also been proposed for improving the quality of foodstuffs, in particular cereal-based foodstuffs; it has, in fact, been noted that they promote dough formation during breadmaking (WONG *et al.*, Cereal Chem. 70, 113-114, 1993), and that, in addition, they decrease the allergenicity of certain foodstuffs.

The inventors have studied the expression of thioredoxin h in cereal grains, in particular in wheat, in order to provide means for controlling this expression.

In the context of these studies, they have isolated a gene, hereinafter named *TaTrxh2*, encoding a soft wheat (*Triticum aestivum*) thioredoxin h, hereinafter named TaTrxh2, the primary structure of which exhibits 97% similarity with that of the soft wheat thioredoxin h TaTrxh1 (GAUTIER *et al.*, 1998, publication mentioned above).

The inventors have also isolated the promoter of the *TaTrxh2* gene and have expressed, in rice, the *gus* reporter gene under the control of this promoter. They thus observed that the expression of the reporter gene was located exclusively in the rice grain and more particularly in the starchy albumen. They also

demonstrated regions involved in the spatial and temporal regulation of this promoter.

The sequence of the *TaTrxh2* gene and of the region in 5′ comprising the promoter are represented in the attached sequence listing under the number SEQ ID NO: 1.

A subject of the present invention is a promoter consisting of a nucleic acid fragment comprising at least one specific functional domain of the promoter of the *TaTrxh2* gene.

The term "promoter" is intended to mean a double-stranded DNA sequence comprising at least the sequences required for initiating the transcription of a gene, optionally combined with sequences for regulating, in *cis*, said transcription; the expression "specific functional domain of a promoter" is intended to mean a sequence of said promoter comprising one or more DNA units involved in transcription initiation, or a double-stranded DNA sequence constituting a regulatory domain comprising one or more of the DNA units involved in the control, in *cis*, of transcription by said promoter.

Promoters in accordance with the invention may, in particular, comprise:

a) the nucleic acid fragment which is represented in the attached sequence listing by the sequence SEQ ID NO: 2, and also on Figure 1, and which corresponds to the 5′ non-coding region of the *TaTrxh2* gene, extending from position -1 to position -1111 relative to the ATG initiation codon, or portions of said fragment, in particular:

* the nucleic acid fragment, the sequence of which extends from position -1 to position -83 relative to the ATG codon of the *TaTrxh2* gene; this fragment comprises the sequences involved in transcription initiation and required for the basic activity of the promoter;

* nucleic acid fragments comprising functional domains involved in regulating the transcription of the *TaTrxh2* gene, and in particular involved in its tissue specificity and/or in its expression at various stages of the development of the plant; it is in particular:

- the nucleic acid fragment, the sequence of which extends from position -591 to position -1111 relative to the ATG codon of the *TaTrxh2* gene; this fragment comprises a regulatory domain involved in inhibiting the expression of the *TaTrxh2* gene in the scutellum epithelium;

- the nucleic acid fragment, the sequence of which extends from position -228 to position -451 relative to the ATG codon of the *TaTrxh2* gene; this fragment comprises a regulatory domain involved in inducing the expression of the *TaTrxh2* gene at the beginning of maturation of the grain;

- the nucleic acid fragment, the sequence of which extends from position -451 to -591 relative to the ATG codon of the *TaTrxh2* gene; this fragment comprises a regulatory domain involved in inducing the expression of the *TaTrxh2* gene in the scutellum epithelium;

- the nucleic acid fragment, the sequence of which extends from position -83 to position -228 relative to the ATG codon of the *TaTrxh2* gene; this fragment comprises sequences involved in inducing the expression in the starch albumen.

b) the nucleic acid fragment constituting the first intron (positions 1232-2203 on the sequence SEQ ID NO: 1) of the *TaTrxh2* gene; this fragment may comprise a regulatory domain of the amplifier type, which increases the level of expression of the *TaTrxh2* gene.

Those skilled in the art may, using the fragments comprising at least one functional domain of the promoter of the *TaTrxh2* gene, specified above, identify more precisely the limits of these functional

domains, and also the DNA units involved in the function of each of them, using techniques which are, in themselves, known, for example using the DNA footprinting technique, incubating these fragments with nuclear extracts of cells from the albumen of the grain, and also with nuclear extracts of cells in which the promoter of the *TaTrxh2* gene is inactive.

The invention in particular encompasses any promoter which may be obtained from a nucleic acid fragment comprising at least one functional domain of the promoter of the *TaTrxh2* gene, using conventional genetic engineering techniques, in particular by mutagenesis and/or genetic recombination. It is thus possible to produce artificial promoters which have the desired level of activity and degree of specificity.

It is thus possible, for example, to inactivate one or more of the functional regulatory domains located in the 5′ non-coding region of the *TaTrxh2* gene, for example by deleting at least one nucleotide or a sequence of nucleotides of the DNA units involved in the function of the domain(s) in question. It is also possible to combine the nucleic acid molecules comprising functional domains of the promoter of the *TaTrxh2* gene with one another, and/or with functional domains originating from promoters other than that of the *TaTrxh2* gene.

According to a preferred embodiment of a promoter in accordance with the invention, it comprises at least the sequences of the promoter of *TaTrxh2* which allow specific expression in the grain, in particular in the starchy albumen.

The invention also encompasses:

- the expression cassettes comprising, besides a promoter in accordance with the invention, a gene of interest placed under the transcriptional control of said

promoter, or a site which allows the insertion of said gene of interest;

- the recombinant vectors resulting from the insertion of a promoter or of an expression cassette in accordance with the invention into a host vector.

The promoters in accordance with the present invention may be used to control the expression of a gene of interest in plant cells, in particular monocotyledon cells.

Said gene of interest may, for example, be either the *TaTrxh2* gene, placed under the control of an artificial promoter, as defined above, derived from the *TaTrxh2* promoter, or a heterologous gene encoding a thioredoxin other than TaTrxh2, or any other protein of interest.

It is possible, for example, to introduce a gene of interest under the control of the promoter of the *TaTrxh2* gene, or of an artificial promoter constructed from the regulatory elements thereof which confer the specificity of expression in the albumen cells, in order to express said gene of interest only in the grain albumen cells. It is also possible to selectively delete the sequences of the promoter of the *TaTrxh2* gene which are responsible for the specificity of expression, in order to construct an artificial promoter which makes it possible to ensure ubiquitous expression of a thioredoxin h, or of another protein of interest.

A subject of the invention is also plant cells and transgenic plants, in particular monocotyledons, and especially cereals, transformed with at least one nucleic acid molecule comprising a promoter in accordance with the invention.

The inventors have thus obtained transgenic rice plants in which a heterologous gene has been placed under the transcriptional control of the promoter of the

TaTrxh2 gene and have observed, in these plants, specific expression in the grain albumen cells.

The transformed cells and the transgenic plants in accordance with the invention can also be used as models to study and/or modify the expression of various genes in the grain albumen cells.

The present invention will be more clearly understood using the further description which follows, which refers to nonlimiting examples illustrating the cloning and characterization of the *TaTrxh2* gene and of its promoter.

EXAMPLE 1:ISOLATION AND CHARACTERIZATION OF THE *TaTrxh2* GENE

1.- Screening a wheat genomic DNA library

A genomic DNA library was prepared from the DNA extracted from leaves of soft wheat (*Triticum aestivum*) of the variety Andain. After partial digestion of the genomic DNA with *Mbo*I, the fragments with a mean size of 15 kb were cloned at the *Bam*HI site of the EMBL3 SP6/T7 phage, which was propagated in the host bacterium K802 -K 802 (*galK2, galT22, HsdR2,* (r_k^-, m_k^+), *mcrA*$^-$, *mcrB*$^-$, *metB1, mrr*$^+$, *supE44*).

6×10^6 clones from the genomic DNA library were plated out and screened with a 669 bp probe (TRX) containing all of the sequence encoding the soft wheat thioredoxin h TaTrxh1 (GAUTIER *et al.*, 1998, publication mentioned above), and the positive clones were then screened by PCR (polymerase chain reaction) using a pair of primers (THP2 and THM2) derived from the same sequence.

One of the clones selected (λ4), which contains an approximately 10 kb wheat genomic DNA fragment, was digested with *PstI*, releasing two fragments, one of 1.5 kb and the other of 3.8 kb, both recognized by the TRX probe. These two fragments were

cloned into the pLITMUS 29 vector (BIOLAB) at the *PstI* restriction site. The two clones obtained are named CTRX3 and CTRX4. The CTRX3 clone corresponds to the 1.5 kb fragment and the CTRX4 clone to the 3.8 kb fragment.

Analysis of the nucleotide sequences of the CTRX4 and CTRX3 clones shows that each contains part of the same gene encoding a wheat thioredoxin h, which is truncated during digestion with *PstI*.

Based on the nucleotide sequences of these two clones, the inventors chose two primers (THP8 and THM8) making it possible to amplify a thioredoxin h gene over a length of approximately 2.6 kb. The PCR was carried out on the undigested DNA of the λ4 clone, and a fragment of the expected size was cloned into the pGEM-T vector (PROMEGA). The clone obtained contained the *TaTrxh2* gene encoding a soft wheat thioredoxin h.

It comprises a 1111 bp promoter region, a 1447 bp coding region and a 131 bp 3′ non-coding region.

The coding region of the *TaTrxh2* gene comprises three exons 120, 123 and 135 bp, separated by two introns of 972 bp and 93 bp. The first exon encodes a 40 amino acid polypeptide, the second exon encodes a 41 amino acid polypeptide containing the active site, and the third exon encodes a 45 amino acid polypeptide.

The nucleotide sequence of the *TaTrxh2* gene encodes a soft wheat thioredoxin h named TaTrxh2, which has 126 amino acids, a calculated molecular mass of 13435 Da and a calculated pI of 5.0.

Comparison of the sequences of the translation products of the *TaTrxh2* gene and of the genes *TaTrxh1* previously described by GAUTIER *et al.* (1998, publication mentioned above) and *TdTrxh1* (hard wheat thioredoxin h) shows that they are very conserved. Specifically, the peptide sequence of TaTrxh2 exhibits 97% similarity and 94% identity with that of TaTrxh1 and 95% similarity and 90% identity with that of TdTrxh1.

The N-terminal domain of TaTrxh2 is shorter than that of TaTrxh1 and TdTrxh1. The primary structure of TaTrxh2 does not contain a signal peptide, suggesting that the protein is located in the cytoplasm. However, it has an N-terminal extension already demonstrated in the primary structure of TaTrxh1 and TdTrxh1, which may correspond to a transmembrane domain. Analysis of the N-terminal extension of TaTrxh2 with the RAO ARGOS program (PC/gene, RAO *et al.*, Biochem. Biophys. Acta 869, 197-214, 1986) reveals a putative transmembrane domain between residues 2 and 19. The active site, made up of the following 5 amino acids: WCGPC, is conserved between the three wheat thioredoxin h TaTrxh2, TaTrxh1 and TdTrxh1.

The sizes of the introns are different from those of the introns of the wheat thioredoxin h genes previously demonstrated by ROBERT (1994), indicating that the *TaTrxh2* gene is different from those genes. The introns of the *TaTrxh2* gene are of the 0 type and are limited in 5′ by the sequence GTA and in 3′ by the sequence CAG, which correspond to consensus sequences of intron-exon limits.

The 3′ non-coding region of the *TaTrxh2* gene has the polyadenylation signal AATAAA common to genes transcribed by RNA polymerase II.

EXAMPLE 2: STRUCTURAL ANALYSIS OF THE PROMOTER OF THE *TaTrxh2* GENE

The promoter of the *TaTrxh2* gene was analysed in order to search for putative regulatory units likely to be involved in controlling expression, and was in particular compared to that of the thioredoxin h genes of *C. reinhardtii* (STEIN *et al.*, Plant Mol. Biol. 28, 487-503, 1995), of tobacco (BRUGIDOU *et al.*, Mol. Gen. Genet 238, 285-293, 1993) and of rice (ISHIWATARI *et al.*, 1995), of the thioredoxin m gene of *C. reinhardtii* (STEIN *et al.*, 1995), and of the murine (MATSUI *et al.*, Gene

152, 165-171, 1995) and human (TONISSEN *et al.*, Gene 102, 221-228, 1992; KAGHAD *et al.*, Gene 140, 6643-6653, 1994) thioredoxins.

The sequence of the promoter of the *TaTrxh2* gene is represented on Figure 1.

The transcription initiation site (represented on Figure 1 in bold and underlined with a double line) is an adenine located at -65 bp from the ATG.

The promoter of the *TaTrxh2* gene contains no consensus sequence corresponding to a TATAn box or to a CAAT box at the expected positions for genes transcribed by RNA polymerase II.

On the other hand, it contains a TATA-like box (AATTTAT, underlined with a double line on Figure 1) at -105 bp from the ATG.

It also contains a GC box (GGGCCGGG, underlined with a dotted line on Figure 1) located at -84 bp from the ATG of the *TaTrxh2* gene. GC boxes are recognized by transcription factors of the Sp1 type (DYNAN *et al.*, Nature 316, 774-778, 1985) and are involved in the constitutive expression of genes. GC boxes are present in all the known promoters of thioredoxin genes.

A sequence rich in adenine, interrupted by a G residue (AAAAAAAGAAAAAAAAA, in bold characters underlined with a single line on Figure 1), is located at -227 bp from the ATG of the *TaTrxh2* gene; sequences of this type have also been identified previously in the promoters of the tobacco and rice thioredoxin h genes.

bHLH sequences (CANNTG), recognized by transcription factors of the helix/loop/helix family, are located at -206 bp and -411 bp from the ATG of the *TaTrxh2* gene; they are represented on Figure 1 in lower case letters.

bzip sequences (ACGT, underlined with a single line on Figure 1), recognized by transcription factors of

the leucine zipper family (bZIP), are located at -251 pb and -184 bp from the ATG of the *TaTrxh2* gene. The bZIP proteins have been described in the activation of expression of genes encoding grain storage proteins. ACGT units have also been described in ABRE (ABA-responsive element) consensus sequences of the promoters of genes, the expression of which is regulated by abscissic acid (ABA) (MUNDY *et al.*, Proc. Natl. Acad. Sci. USA, 87, 1406-1410, 1990).

Two pyrimidine boxes (CCTTTCTCT and TCTTTCTTC, boxed on Figure 1) are, respectively, located at -553 bp and -541 bp from the ATG of the *TaTrxh2* gene. Pyrimidine boxes (CCTTTT) are involved in the regulation of expression by giberellic acid, generally in combination with GARE (GA-responsive element) sequences (TAACAAA) (HUANG *et al.*, Plant Mol. Biol. 14, 115-121, 1990), and O2S (opaque-2-binding sequence) or I box (TATCCAT) sequences (GUBLER *et al.*, Plant Cell 4, 1435-1441, 1992; LANAHAN *et al.*, Plant Cell, 4, 203-211, 1992), with which they organize into a complex named GARC (GA-responsive complex) (BETHKE *et al.*, Bot. 48, 1337-1356, 1997).

No GARE or O2S sequence was revealed in the 1111 bp upstream of the ATG of the *TaTrxh2* gene.

A TGTGTGAGCA unit (in bold characters and underlined with a dotted line on Figure 1) is located at -403 bp from the ATG of the *TaTrxh2* gene. This unit differs only by the presence of an additional G residue, from the GCN4-like consensus sequence (TGTGTGACA) of the "albumen box" involved in the albumen-specific expression of wheat glutenin genes (HAMMOND-KOSACK *et al.*, EMBO J. 12, 545-554, 1993). Hovever, the other unit of the albumen box, named EM (TGTAAAAGT), and the presence of which is also required for albumen-specific expression (ALBANI *et al.*, Plant Cell 9, 171-184, 1997), was not revealed in the 1111 bp upstream of the ATG of the *TaTrxh2* gene.

Trimeric diades CAA and TTG (in italics on Figure 1) separated by 10 bases are present, respectively, at -107 bp and -97 bp from the ATG of the *TaTrxh2* gene. These units are associated with specific expression in the aleurone layer (THOMAS *et al.*, Plant Cell 2, 1171-1180, 1990).

EXAMPLE 3: FUNCTIONAL ANALYSIS OF THE PROMOTER OF THE *TaTrxh2* GENE.

The 1111 bp sequence in 5' of the ATG of the *TaTrxh2* gene, or various fragments of this sequence, were cloned upstream of the coding sequence of the *gus* reporter gene in the pSPORT1-GUS vector. The pSPORT1-GUS vector (DIGEON, 1997) contains the coding sequence of the *gus* (*E. coli* β-glucuronidase) gene and the *nos-ter* terminator of the nopaline synthase gene, inserted at the EcoRI-HindIII site of the pSPORT1 vector (GIBCO BRL).

The constructs prepared are as follows:

- P1: this construct comprises all of the 1111 bp sequence upstream of the ATG of the *TaTrxh2* gene;
- P2: this construct comprises the 589 bp sequence upstream of the ATG of the *TaTrxh2* gene;
- P3: this construct comprises the 481 bp sequence upstream of the ATG of the *TaTrxh2* gene;
- P4: this construct comprises the 228 bp sequence upstream of the ATG of the *TaTrxh2* gene;
- P5: this construct comprises the 83 bp sequence upstream of the ATG of the *TaTrxh2* gene;

The limits of the regions of the promoter of the *TaTrxh2* gene which were used in the constructs are indicated on Figure 1.

As a positive control, the pUGC1 vector (CHAÏR *et al.*, 1996), which allows constitutive and ubiquitous expression of the *gus* gene under the control of the promoter, of the first exon and of the first intron of the gene encoding maize ubiquitin was used.

The pSPORT1-GUS vector was used as a negative control.

These various constructs were transferred, by bombardment according to the protocol described by FAUQUET *et al.* (Proc. Third. Int. Rice Genet. Symp., Ed. G.S. Khush, 153-165, 1996), into young embryogenic calluses of rice (var. japonica IRAT 349) which derive from the proliferation of the scutellum of the mature embryo. All the constructs tested were cotransferred with the pILTAB227 vector (FAUQUET *et al.*, 1997), which confers hygromycin resistance and which allows selection of the transformed cells.

A mixture of the vector carrying the construct to be tested and of the pILTAB227 vector (vector to be tested/pILTAB227 molar ratio = 4/1) is used to coat gold microparticles, in a proportion of 5 µg of total DNA (3 µg of DNA to be tested + 2 µg of pILTAB227) at a concentration of 1 µg/µl, per 3 mg of a mixture, in equal amounts, of gold microparticles having diameters of 1.0 µm and 1.6 µm, in suspension in 50 µl of distilled water.

The bombardment is carried out using a PDS-1000/He particle gun (PARTICLE DELIVERY SYSTEM, BIORAD).

The bombarded embryogenic calluses are then screened on a selection medium containing hygromycin. The hygromycin-resistant calluses are selected and placed on hygromycin-free regeneration medium. The regenerated plants (F0 generation) are then transferred into pots and, after acclimatization in a phytotron, are cultivated under glass.

The expression of the *gus* gene was sought in the vegetative organs and in the grains of the rice plants of the T0 and T1 generations. Only the plants which were fertile and had a normal phenotype were selected for analysis. Integration of the transgene into

the plants analysed was verified by PCR and Southern transfer.

The β-glucuronidase activity was detected using a histochemical test, detecting the blue coloration resulting from hydrolysis of 5-bromo-4-chloro-3-indolyl β-D-glucuronic acid (X-GLU), and was quantified using a fluorometric test, measuring the 4-methylumbelliferone (MU) formed from 4-methylumbelliferyl β-glucuronic acid, according to the protocol described by JEFFERSON *et al.* (Plant Mol. Biol. Report 5, 387-405, 1987).

1. Expression of the *gus* gene in the vegetative organs

The analysis was carried out on the roots, the culms and the leaves.

In the case of the nontransformed rice plants, or of those transformed with the pSPORT1-GUS vector, the histochemical analysis reveals no GUS activity, and the fluorimetric measurements reveal only very low, or even zero, activity.

In the case of the rice plants transformed with the pUGC1 vector, high GUS activity (greater than 500 pmol MU/min/mg of protein) is observed in all the vegetative organs tested.

In the case of the rice plants transformed with the P1, P2, P3, P4 or P5 constructs, no coloration is observed in the vegetative organs incubated in the presence of X-GLU, and the GUS activity measured by fluorometry is not significantly different from those measured for the plants transformed with the pSPORT1-GUS vector or the nontransformed plants. These results show that the promoter of the *TaTrxh2* gene does not allow expression of the *gus* gene in the vegetative organs.

2. Expression of the *gus* gene in the grains

Histochemical analysis

In whole grains

The rice grains, taken at 35 DAF (days after fertilization) were cut in the longitudinal direction and then incubated in the presence of X-GLU.

No coloration is detected in the rice grains transformed with the pSPORT1-GUS vector.

An intense coloration of the entire grain is, on the contrary, detected for the rice grains transformed with the pUGC1 vector.

In the case of the rice grains transformed with the P1, P2, P3 or P4 constructs, a blue coloration is detected in the albumen of the grains, but not in the embryo (cotyledonary axis and scutellum), the envelopes or the spikelets. In the albumen, this coloration appears in particular at the periphery of the embryo, above the scutellum epithelium, and in a median zone of the albumen over the entire length of the grain.

This coloration is less intense and appears less rapidly than that observed in the rice grains transformed with pUGC1. The intensity of the coloration appears to vary depending on the construct (P1, P2, P3 or P4) concerned. The strongest intensity is observed in the rice grains transformed with the P2 construct and the weakest intensity is observed in those transformed with the P4 construct. These results are confirmed by the analysis of the T1 grains, in which the coloration appears more rapidly than in the T0 grains and is more intense.

In the case of the rice grains transformed with the P5 construct, no coloration is detected, either in the embryo, in the albumen or in the envelopes.

In the various tissues of the grain

In order to accurately determine the location of the expression of the *gus* gene under the control of the promoter of the *TaTrxh2* gene, histological sections were prepared and observed using a photon microscope.

The observations show that, for the P1, P2, P3 or P4 constructs, the labelling is located in a small number of cells of the starchy albumen, located at the periphery of the embryo and in the central zone of the starchy albumen. The cells of the embryo, of the aleurone layer and of the envelopes are not labelled. For the P5 construct, no labelling is visible.

Fluorimetric analysis

The GUS activity was measured, firstly, on the embryos and, secondly, on the albumen of the rice grains.

The GUS activity is zero or very low in the rice grains transformed with the pSPORT1-GUS vector, as in the nontransformed rice grains. On the other hand, it is very high (>500 pmol/MU/min/mg of protein) in the embryo and the albumen of the rice grains transformed with the pUGC1 vector.

The GUS activity measured in the embryos of the rice grains transformed with the P1, P2 P3, P4 or P5 constructs is not significantly different from that measured in the nontransformed rice grains or rice grains transformed with the pSPORT1-GUS vector.

On the other hand, the activity measured in the albumen of the rice grains transformed with the P1, P2, P3 or P4 constructs is 25 to 40 times higher than that measured in the albumen of the nontransformed rice grains or the rice grains transformed with the pSPORT1-GUS vector. It ranges from 40 pmol/MU/min/mg of protein, for the rice grains transformed with the P2 construct, to 25 pmol/MU/min/mg of protein, for those transformed with the P4 construct. For the P5 construct, no GUS activity is detected in the grains.

These results show that the region (-1111 bp to -83 pb) of the promoter of the *TaTrxh2* gene allows expression of the *gus* gene only in the cells of the starchy albumen, and that only the deletion which leaves only 83 bp upstream of the ATG suppressed the sequences responsible for the spatial expression, some of which are probably located in the region of the promoter which is between -228 bp and -83 pb.

The GCN4-like unit identified during the structrual analysis of the promoter of the *TaTrxh2* gene is therefore apparently not the only one responsible for the tissue specificity of the expression; specifically, despite being deleted in the P3 and P4 constructs, the expression of the *gus* gene remains specific for the albumen of the grain.

Two sequences: AACAAATCC and AACAAAGTG (represented in bold characters on Figure 1), are present at -51 bp and -381 bp, respectively, relative to the ATG of the *TaTrxh2* gene. These sequences exhibit similarity with AACA units (AACAAACTCTATC) recently demonstrated in the promoters of 6 genes encoding rice glutelins, and involved in the albumen-specific expression of these genes.

EXAMPLE 4: EVOLUTION OF THE EXPRESSION OF THE *GUS* GENE DURING THE DEVELOPMENT OF THE GRAINS OF THE TRANSGENIC RICE PLANTS

The expression of the *gus* gene was monitored during the maturation and germination of rice grains transformed with the P1, P2, P3, P4 or P5 constructs.

1. During maturation

Three stages were analysed: 10 DAF, 25 DAF and 35 DAF. The expression of the *gus* gene was evaluated either by histochemical location of the GUS activity or by detection of the transcripts by Northern transfer.

GUS activity

The histochemical analysis shows that, for the three stages of maturation studied, 10, 25, and 35 DAF, GUS activity is always detected in the starchy albumen of the rice grains transformed with the P1, P2, P3 or P4 constructs. On the other hand, for the P5 construct, no GUS activity is detected.

At 10 DAF, the GUS activity is detected in the starchy albumen, at the periphery of the embryo. At 25 DAF, the GUS activity progresses towards the median zone of the starchy albumen. At 35 DAF, the GUS activity is detected over the entire surface of the starchy albumen.

The intensity of the coloration varies with the nature of the construct and the stage of maturation, in particular in the case of the P4 construct, for which the coloration is very difficult to detect at the start of maturation.

These results are confirmed by the more detailed observations in each tissue of the grain, which show that:

- At 10 DAF: for the P1, P2 or P3 constructs, the GUS activity is very high in the cells of the starchy albumen at the periphery of the embryo, and it is not detected in the cells of the median zone of the starchy albumen. For the P4 construct, the GUS activity in the cells of the starchy albumen is very low, or even undetectable. In addition, in the rice grains transformed with the P2 construct, GUS activity is also detected in the cells of the scutellum epithelium.

- At 25 DAF: for the P1, P2 or P3 constructs, the GUS activity decreases in the cells of the starchy albumen at the periphery of the embryo and increases in those of the central zone of the starchy albumen; for the rice grains transformed with the P2 construct, GUS activity is no longer detected in the cells of the

scutellum epithelium. In the rice grains transformed with the P4 construct, the GUS activity increases in the cells of the starchy albumen which are located at the periphery of the embryo and in the central zone.

- At 35 DAF: the GUS activity is much lower than at 25 DAF in all the cells of the starchy albumen of the rice grains transformed with P1, P2 or P3; in the rice grains transformed with the P4 construct, the GUS activity is detected in the cells of the median zone of the starchy albumen.

For the P5 construct, no GUS activity is detected, whatever the stage of maturation or the tissue of the grain analysed.

Whatever the construct used, no GUS activity is detected, during maturation, in the cells of the embryonic axis, of the aleurone layer or of the envelopes of the grains.

Detection of the *gus* gene transcripts

The presence of the *gus* gene transcripts was sought in the total RNAs extracted, at the various stages of maturation, from the rice grains transformed with the P1, P2, P3, P4 or P5 constructs. The detection was carried out by Northern transfer, using the P3+GUS probe. This 2.6 kb probe comprises the 481 bp sequence upstream of the ATG of the *TaTrxh2* gene, the coding sequence of the *gus* gene and the terminator of the *nos* gene.

For each of the P1, P2, P3, P4 or P5 constructs, this probe makes it possible to detect the presence of transcripts which have an expected size of between 1.9 and 2.4 kb, depending on the construct.

The presence of these transcripts varies depending on the construct used for the transformation and on the stage of maturation.

For the rice plants transformed with the P1 construct, the transcripts are detected at the 3 stages of maturation, with a maximum at mid-maturation.

For the rice plants transformed with the P2 construct, the transcripts are detected at the start of maturation.

For the rice plants transformed with the P3 construct, the transcripts are detected at the start and at the end of maturation of the grain.

For the rice plants transformed with the P4 construct, the transcripts are detected at mid-maturation and at the end of maturation of the grain.

For the rice plants transformed with the P5 construct, no *gus* gene transcript is detected, whatever the stage of maturation analysed.

2. During germination

For the study of the expression of the *gus* gene during germination, the GUS activity was analysed by histochemistry in rice grains transformed with the P1, P2, P3, P4 or P5 constructs.

For each construct, 10 grains were left to germinate in the dark and taken at various times after soaking: 0, 12, 24, 48 and 72 hours.

The GUS activity is detected in the starchy albumen of the rice grains transformed with the P1, P2, P3 or P4 constructs, whatever the stage of germination. On the other hand, for the rice grains transformed with the P5 construct, no GUS activity is detected.

The study of the accumulation of the *gus* gene transcripts in the grains transformed with the P1, P2, P3 or P4 constructs shows that these transcripts are not accumulated during germination, regardless of the construct.

These results indicate that the promoter (1111 bp upstream of the ATG) of the *TaTrxh2* gene does not allow expression of the *gus* gene during germination. The GUS activity detected in the germinated grains is certainly a residual activity due to the very great stability of β-glucoronidase in the grain.

Conclusion

The analysis of the expression of the *gus* gene under the control of the promoter of the *TaTrxh2* gene, during the development of the rice grains transformed with the P1, P2, P3, P4 or P5 constructs demonstrates an effect of the deletions of the promoter of the *TaTrxh2* gene on the temporal and spatial expression of the *gus* gene in the grains of the rice plants transformed.

The P1, P2 and P3 constructs allow earlier expression of the *gus* gene than the P4 construct, during maturation. The P5 construct does not allow expression of the *gus* gene since no transcript is detected. In fact, *gus* gene transcripts are detected at 10 DAF for the 3 constructs P1, P2 and P3, and only 25 DAF for the P4 construct. This suggests that the differences in level of expression previously noted between the P2 and P4 constructs are probably the result of a delay in the expression of the *gus* gene under the control of the P4 promoter, rather than the result of a lower level of expression. The region of the promoter of the *TaTrxh2* gene which is between -1111 bp and -228 bp definitely contains regulatory sequences which allow expression of the *gus* gene in the first stages of maturation.

With regard to the spatial expression, the region of the promoter of the *TaTrxh2* gene which is between -1111 bp and -591 bp probably contains a sequence which inhibits the expression of the gene in the scutellum epithelium. In fact, when it is deleted (P2 construct), expression of the *gus* gene is observed in this tissue. Conversely, the region between -591 bp and -451 bp is thought to contain a sequence which activates the expression in the scutellum epithelium since, when it is deleted (P3 construct), there is no longer an expression of the *gus* gene in this tissue.

The results show that, during the maturation of the rice grains, the promoter of the *TaTrxh2* gene

allows expression of the *gus* gene, specific to the starchy albumen. This expression is detected in a restricted number of cells distributed in a central zone of the albumen and at the periphery of the embryo.